

A Data-Driven Sound Analysis Approach for Detecting Mechanical and Electrical Anomalies in Motor Systems

T. Rizwana¹, K. Srinivas^{2*}, Perugu Charmi³, Malyala Saivaran³, Kotthapalli Ramya³, Mohd Usman Ahmed³

¹Assistant Professor, ²Associate Professor & Head, ³UG Student, ^{1,2,3}Department of Computer Science and Engineering (Data science)

^{1,2,3}Vaagdevi Engineering College, Bollikunta, Warangal, 506005, Telangana, India

*Correspondence: K. Srinivas (srinivas_k@vecw.edu.in)

Abstract

Industrial electric motors are fundamental to modern manufacturing operations, making accurate and reliable fault diagnosis crucial for minimizing downtime and maintenance costs. Acoustic monitoring has gained attention as a non-intrusive alternative to traditional condition monitoring methods such as vibration and thermal analysis. However, widely used handcrafted features like Mel Frequency Cepstral Coefficients (MFCCs) and spectral centroids often fail to effectively represent the complex temporal and spectral characteristics of industrial audio signals. These conventional features are highly sensitive to environmental noise and variations in machine operating conditions, which limits their ability to generalize across diverse fault types including bearing faults, gearbox defects, and fan imbalances. To overcome these challenges, this work introduces a robust diagnostic framework based on Waveform-based Language Model (WavLM) transformer embeddings. Unlike traditional approaches, WavLM utilizes self-supervised learning to extract high-dimensional, context-aware representations directly from raw audio data, offering improved resilience to noise and enhanced feature richness. These embeddings are then utilized as inputs to multiple machine learning classifiers, including Categorical Boosting (CB) and Decision Tree (DT) models, for fault classification. Experimental evaluation demonstrates that the proposed Histogram-Based Gradient Boosting (HGB) classifier achieves superior performance, attaining an accuracy of 95%, significantly outperforming conventional machine learning approaches. Additionally, a user-friendly graphical user interface (GUI) is developed to support efficient data handling, model training, and real-time prediction. This study demonstrates the effectiveness of transformer-based audio features for scalable and reliable predictive maintenance.

Keywords: WavLM (Transformer Embeddings), Acoustic Monitoring, Predictive Maintenance, Industrial Electric Motors, Self-Supervised Learning.

1. Introduction

Electric induction motors are among the most essential components driving modern industrial production and daily life. They are widely employed across manufacturing industries as well as in domestic applications due to their efficiency and reliability. An electric motor converts electrical energy into mechanical energy by generating torque through the interaction between magnetic fields and electric current within its windings [1]. Any failure or unexpected shutdown of such critical machinery can lead to severe consequences, including equipment damage, financial losses, operational disruptions, environmental hazards, and even risks to human safety. Therefore, the development of effective motor fault diagnosis techniques is of paramount importance [2]. Fault diagnosis systems enable early detection of motor abnormalities, allowing timely maintenance and reducing repair costs while improving operational efficiency and minimizing production downtime, as shown in figure. 1. Traditional diagnostic methods rely heavily on manually extracted features from time-domain,

frequency-domain, and time–frequency analyses [3]. However, these approaches are often complex, time-consuming, and prone to uncertainty due to their dependence on handcrafted feature engineering.

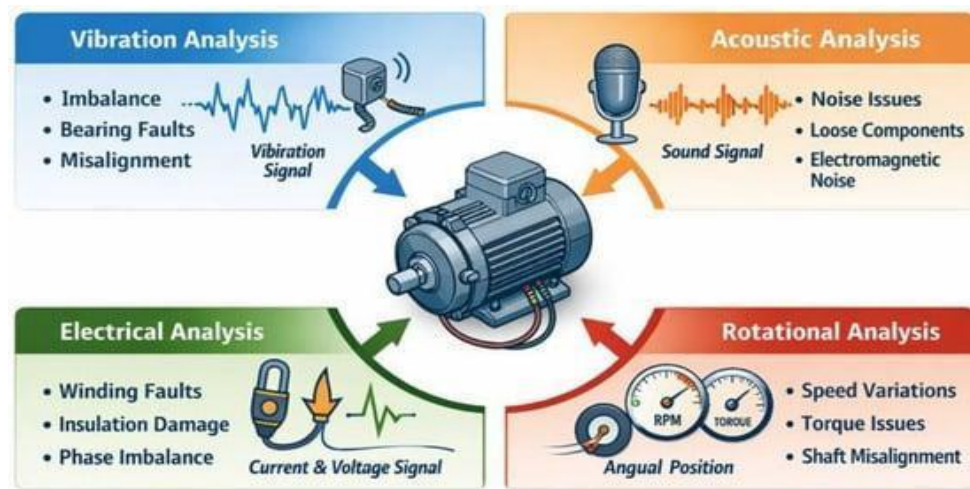


Figure. 1: Fault diagnosis system on electric motor.

With the rapid advancement of industrial systems, motor-generated data has evolved into a form of “big data,” characterized by high volume, diversity, velocity, and low value density [4]. Conventional methods struggle to handle such complex data efficiently. The emergence of artificial intelligence has significantly transformed fault diagnosis, shifting it from traditional techniques to intelligent, data-driven approaches. Although early artificial neural networks offered adaptive feature learning capabilities without explicit mathematical modelling, they suffer from limitations such as overfitting, vanishing gradients, local minima issues, and reliance on prior knowledge, which restrict their diagnostic performance [5].

2. Literature Survey

Yang, et al. [6] covered four traditional types of deep learning models: deep belief networks (DBN), autoencoders (AE), convolutional neural networks (CNN), and recurrent neural networks (RNN), and highlights their use in detecting faults of electric motors. Finally, the issues and obstacles that deep learning encounters in the fault detection mechanism as well as the prospects are discussed and summarized. Deep learning in the fault detection of electric equipment has shown comparatively better results than traditional approaches because of its more powerful and sophisticated feature extraction capabilities. Electric motors are used extensively in numerous industries, and their failure can result not only in machine damage but also a slew of other issues, such as financial loss, injuries, etc. As a result, there is a significant scope to use robust fault diagnosis technology.

Gu, et al. [7] proposed a model for motor fault diagnosis based on deep learning and multi-sensor information fusion. Firstly, a correlation adaptive weighting method is proposed in this paper, and it is used to integrate the collected multi-source homogeneous sensor information into multi-source heterogeneous sensor information through the data layer fusion. Secondly, the 1D-CNN is used to carry out feature extraction, feature layer fusion, and fault classification of multi-source heterogeneous information of the motor. Finally, the data of seven states (one healthy and six faulty) of the motor are collected by the motor drive test bench to realize the model’s training, testing, and verification. The experimental results show that the fault diagnosis accuracy of the model is 99.3%.

Khaneghah, et al. [8] aimed to provide a comprehensive overview of potential faults in EV motor drives and battery systems, while also reviewing the latest state-of-the-art research in EV fault detection. The information presented herein can serve as a valuable reference for future endeavours in this field. In the realm of EV applications, Permanent Magnet Synchronous Motors (PMSMs) and lithium-ion battery packs have garnered significant attention. Consequently, fault detection methods for PMSMs and their drives, as well as for lithium-ion battery packs, have become a prominent area of research. An effective FDD approach must possess qualities such as accuracy, speed, sensitivity, and cost-effectiveness. Traditional FDD techniques include model-based and signal-based methods. However, data-driven approaches, including machine learning-based methods, have recently gained traction due to their promising capabilities in fault detection.

Choi, et al. [9] developed a Convolutional Autoencoder (CAE) and deep neural network (DNN)-based model optimized for real-time signal processing and high accuracy in motor fault diagnosis. This model learns complex patterns from voltage and current data and precisely analysed them in combination with DNN through latent space representation. Traditional diagnostic methods relied on vibration and current sensors, empirical knowledge, or harmonic and threshold-based monitoring, but they had limitations in recognizing complex patterns and providing accurate diagnoses. Our model significantly enhances the accuracy of power data analysis and fault diagnosis by mapping each phase (R, S, and T) of the electrical system to the red, green, and blue (RGB) channels of image processing and applying various signal processing techniques. Optimized for real-time data streaming, this model demonstrated high practicality and effectiveness in an actual industrial environment, achieving 99.9% accuracy, 99.8% recall, and 99.9% precision.

Li, et al. [10] detected the equipment status in extreme environments. However, conventional Infrared thermal images inevitably show a large amount of noise interference, which affects the analysis results. In addition, each motor may only possess a small amount of fault data in practice, as collecting an infinite amount of motor data to train the diagnostic system is impossible. To overcome these problems, a novel automatic fault diagnosis system is proposed in this study. Data features are enhanced by a normalization module based on color bars first, as the same color in various infrared thermal images represent different temperatures. Then, the few-shot learning method is used to diagnose the faults of unseen electric motors. In the few-shot learning method, the minimum dataset size required to expand system universality is fifteen pieces, effectively solving the universality problem of artificial-to-natural data migration. The method saves a large amount of training data resources and the experimental training data collection. The accuracy of the fault diagnosis system achieved 98.9% on similar motor datasets and 91.8% on the dataset of motors that varied a lot from the training motor.

Huang, et al. [11] proposed a framework using a fine-tuned Qwen2.5-7B model for EV motor fault diagnosis and predictive maintenance. It merges current and vibration signals, extracting features via time-frequency analysis to assess the health state of motor. Lightweight fine-tuning with LoRA and QLoRA reduces training costs and improves generalization and real-time capabilities. Experiments show a 96.5% accuracy on single datasets, 91.2% across conditions, and 82.7% with small samples, outperforming CNN and LSTM models. This analysis not only highlights key features like vibration spectral energy and current fluctuation amplitude, but also enhances the interpretability of results. This study introduces a novel approach to EV motor fault diagnosis, demonstrating the great potential of large language models to tackle fault diagnosis of EV motor.

Tao, et al. [12] provided a comprehensive review of fault diagnosis techniques for IWMs. First, typical faults in IWMs are analyzed with a focus on their unique structural and failure characteristics. Then,

the applications and recent research progress of three major categories of fault diagnosis approaches model-based, signal-based, and knowledge-based methods in the context of IWMs are critically reviewed. Finally, key challenges and pain points in IWM diagnosis are discussed, along with promising future research directions. Although considerable progress has been made in fault diagnosis techniques related to IWMs, a systematic review in this area is still lacking.

Kumar, et al. [13] proposed a recurrent convolutional neural network (RCNN) for effective defect detection in motors, taking advantage of the advances in deep learning techniques. The proposed approach applies long short-term memory (LSTM) layers to capture the temporal dynamics essential for fault detection and convolutional neural network layers to mine local features from the segmented vibration data. This hybrid method helps the model to learn complicated representations and correlations within the data, leading to improved fault detection. Model development and testing are conducted using a sizable dataset that includes various kinds of motor defects under differing operational scenarios. The results demonstrate that, in terms of fault detection accuracy, the proposed RCNN-based strategy performs better than the traditional fault detection techniques. The performance of the model is assessed under varying vibration data noise levels to further guarantee its effectiveness in practical applications.

Sobhi, et al. [14] assigned to a single electrical circuit, meaning a failure in one could damage other motors on that circuit. There is thus a need for condition monitoring of aggregations of small motors. They report on ongoing research to develop a machine-learning-based solution for fault detection in multiple small electric motors. Shallow and deep learning approaches to this problem are investigated and compared, with a hybrid deep/shallow system ultimately being the most effective. Such capabilities are usually not cost-effective for small (under ten horsepower) motors, as they are inexpensive to replace. However, large industrial sites may use hundreds of these small motors, often to drive cooling fans or lubrication pumps for larger machines.

Soresini, et al. [15] focussed on durability testing of Permanent Magnet Synchronous Motors for automotive applications, using Autoencoders (AEs) to predict and prevent failures. This AI-based fault detection strategy employs acceleration signals coming from electric motors tested under challenging conditions with significant variations in torque and speed. This approach goes beyond typical fault detection in steady-state conditions. Based on a review of Neural Networks, including Variational Autoencoders (VAEs), Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, the performance of six AI architectures is compared: AE, VAE, 1D CNN AE, 1D CNN VAE, LSTM AE, and LSTM VAE. The 1D CNN AE outperformed the other networks in fault detection, showing high accuracy, stability, and computational efficiency. The model is integrated into an algorithm for semi-real-time fault monitoring. The algorithm effectively detects potential motor failures in real-world scenarios, including bearing faults, mechanical misalignments, and progressive wear of components, thereby proactively preventing damage and halving test bench downtime.

3. Proposed Methodology

The proposed methodology presents a structured and data-driven framework for fault diagnosis in electric motors using acoustic signal analysis. The system is designed to detect and classify motor faults by leveraging deep learning-based feature extraction and machine learning classification techniques. The analytical pipeline begins with audio data acquisition and progresses through preprocessing, feature extraction, model training, evaluation, and real-time prediction stages. The overall architecture of the framework is illustrated in Figure. 2. By integrating transformer-based audio representation (WavLM) with multiple machine learning classifiers, the system effectively captures complex temporal and

spectral patterns present in motor sounds. The framework is capable of handling noisy, real-world industrial audio data and ensures robust performance across various fault conditions. Comparative evaluation across multiple models enhances reliability and consistency. The system supports both offline model training and real-time prediction through an interactive GUI.

User Interface (Audio Input / Dataset Upload)

- Users upload motor sound datasets organized into fault-specific categories or provide individual audio samples for prediction.
- The interface enables dataset selection, model training, and real-time fault diagnosis.
- All user interactions are forwarded to the backend processing system.

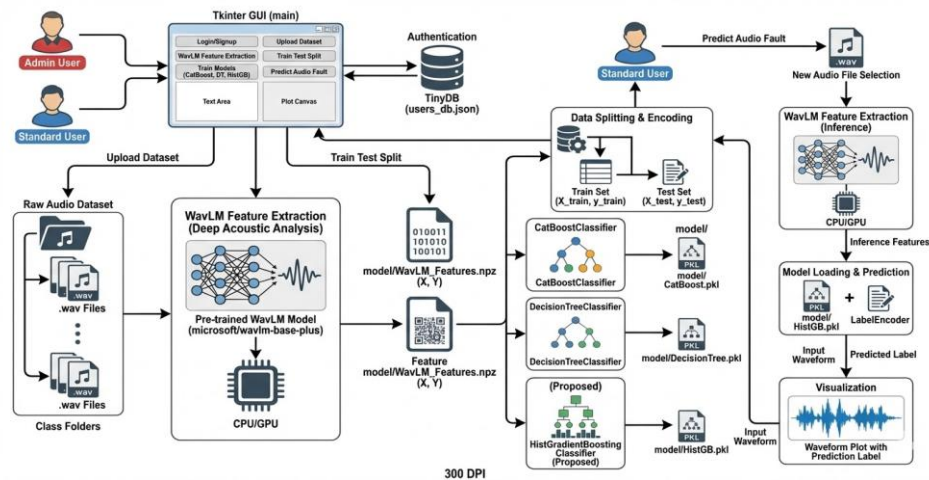


Figure. 2: System overview architecture.

Backend Processing Environment

- Controls the workflow across feature extraction, model training, and prediction stages.
- Manages model storage, loading, and execution.
- Ensures seamless communication between input modules and output visualization.

Raw Acoustic Data and Preprocessing

- **Raw Acoustic Data:** Consists of motor sound recordings in WAV format, including variations caused by noise, load changes, and environmental disturbances.
- **Audio Preprocessing Module:** Loads audio signals and standardizes the sampling rate to **16 kHz**. It normalizes signal amplitude, removes inconsistencies, and prepares audio data for deep feature extraction.

Feature Extraction and Management

- **Feature Extraction using WavLM:** Utilizes a pretrained transformer-based WavLM model to extract high-dimensional, context-aware embeddings. It applies mean pooling to generate fixed-length feature vectors, capturing both temporal and spectral characteristics effectively.
- **Feature Storage and Management:** Stores extracted features in compressed format (.npz) to enable faster reuse without repeated computation, improving system efficiency and scalability.

Data Preparation and Machine Learning Models

- **Data Preparation:** Encodes class labels using Label Encoding and splits the dataset into training and testing sets with stratification to ensure a balanced representation of fault classes.
- **Machine Learning Models:** Implements HGB as the proposed model, with CB Classifier and DT Classifier serving as baseline models for comparative analysis. Each model independently learns fault patterns from the extracted features.

Prediction, Evaluation, and GUI

- **Prediction and Fault Classification:** Processes new audio inputs through the same pipeline, using the trained HGB model to convert predicted labels into human-readable fault categories.
- **Performance Evaluation:** Computes accuracy, precision, recall, and F1-score. Generates confusion matrices, ROC curves, and AUC for multi-class evaluation and comparative analysis.
- **GUI and User Management:** Developed using Tkinter for interactive control, supporting role-based access (Admin/User). Includes secure authentication using Tiny DB and hashed credentials.

System Output and Interpretation

- Displays the predicted fault type along with a waveform visualization.
- Provides clear and interpretable diagnostic results to support critical decision-making for predictive maintenance.

4. Results description

The results of the proposed system demonstrate the effectiveness of transformer-based acoustic feature extraction combined with machine learning models for motor fault diagnosis. The WavLM-based embeddings successfully capture complex temporal and spectral characteristics of motor sounds, leading to improved classification performance. Among the evaluated models, the HGB classifier achieved the highest accuracy, outperforming CB and DT models. The system also showed strong performance across precision, recall, and F1-score metrics, indicating reliable fault detection. Confusion matrix and ROC curve analysis further validate the robustness and generalization capability of the model. Additionally, the GUI-based implementation enables real-time prediction with clear visualization of results.

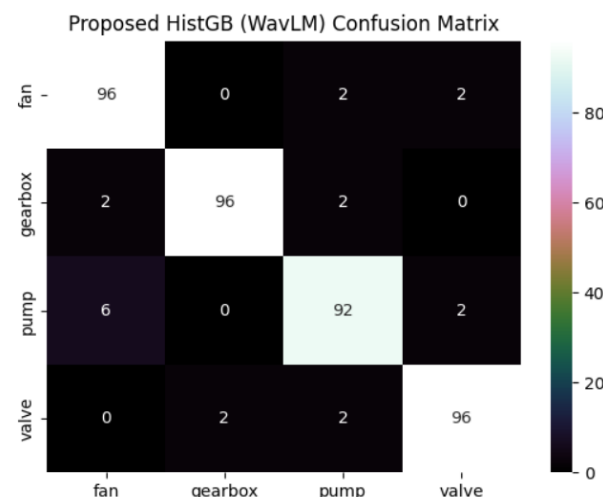


Figure. 3: HGB Classifier obtained for Confusion Matrix.

Figure 3 presents the confusion matrix of the proposed HGB classifier, which achieves the best performance among all tested models. The model demonstrates near-perfect classification accuracy, with almost all samples correctly mapped to their respective classes. The strong diagonal presence and minimal off-diagonal values confirm that HGB effectively leverages Wav LM embeddings, capturing subtle acoustic variations and providing highly reliable fault predictions for electric motor systems.

Figure 4 shows ROC curve for the proposed HGB (WavLM) model demonstrates perfect classification performance across all fault classes, with fan, gearbox, pump, and valve each achieving an AUC of 1.00. All curves tightly follow the top-left boundary and remain far above the diagonal reference line, indicating an optimal balance of maximum true-positive rates and near-zero false positives. This result highlights the superior discriminative capability and robustness of the proposed HGB model when combined with WavLM acoustic features.

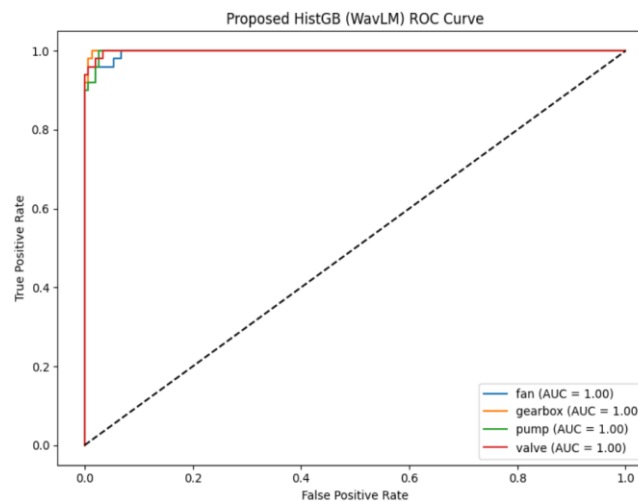


Figure. 4: ROC obtained for proposed HGB

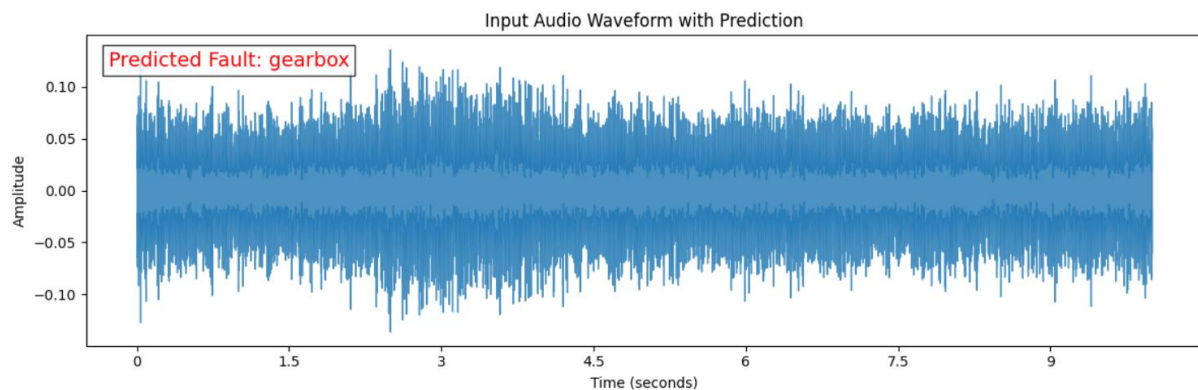


Figure. 5: Waveform Output – Predicted Label: Gearbox.

Figure 5 displays the waveform corresponding to an audio sample classified as fan. The amplitude variations match the typical behaviour of fan noise—more uniform, with regular low-frequency oscillations. The correct classification further validates the model's discriminative power in distinguishing between subtle acoustic signatures found in similar industrial components

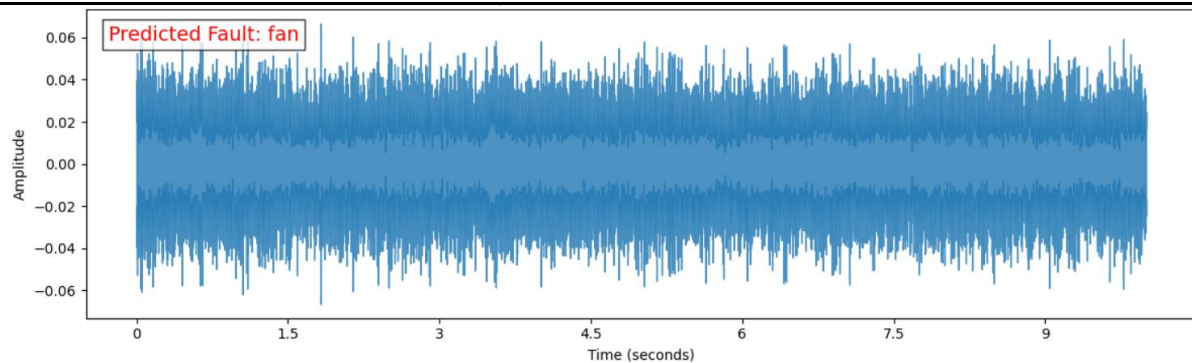


Figure. 6: Waveform Output – Predicted Label: Fan.

Figure. 6 displays the waveform of another audio sample that has been classified as fan by the proposed acoustic fault-diagnosis system. The waveform demonstrates a highly consistent amplitude distribution throughout the entire 10-second duration, which is characteristic of continuous fan operation. The signal exhibits uniform oscillations with low-to-medium intensity variations, reflecting the steady airflow and mechanical rotation typically produced by industrial fans. This visual pattern aligns closely with the acoustic signatures learned during model training, confirming the system's ability to accurately recognize recurring fan noise. The clear time-domain representation also helps confirm that the audio recording is free of clipping or distortions, ensuring reliability of the prediction. The consistency across multiple correctly classified waveforms further validates the robustness of the WavLM-based feature extraction pipeline and the strength of the proposed classifier in differentiating between subtle machine sound profiles.

Table. 1: Overall comparison table

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CB	75.50	75.81	75.50	75.26
DT	73.75	77.24	73.75	73.46
HGB	95.00	95.02	95.00	94.99

The performance of the three machine learning models CB, DT, and HGB is evaluated using key metrics such as accuracy, precision, recall, and F1-score. Among these models, HGB demonstrates the highest effectiveness with an accuracy of 95.00%, indicating its superior ability to correctly classify instances. In comparison, CB achieves an accuracy of 75.50%, showing moderate performance across all evaluation metrics. The DT model records a slightly lower accuracy of 73.75%, despite having relatively higher precision than CB. HGB consistently outperforms the other models in precision (95.02%), recall (95.00%), and F1-score (94.99%). These results highlight the robustness and reliability of HGB for this classification task. Hence, HGB can be considered the most accurate and well-balanced model among the three.

5. Conclusion

The proposed system for Electric Motor Fault Diagnosis using Wav LM Embeddings and Machine Learning successfully demonstrates a highly accurate and efficient methodology for identifying various industrial machine faults such as fan, gearbox, pump, and valve anomalies. Traditional spectral features were replaced with advanced Wav LM transformer-based embeddings, which capture deeper temporal,

harmonic, and structural acoustic characteristics. These embeddings significantly improved the learning capability of all evaluated models. Among the three classifiers such as CB, DT, and the Proposed HGB the HGB classifier achieved the highest performance, reaching 95% accuracy, a major improvement over CB (75.50%) and DT (73.75%). This improvement clearly validates that deep audio representations combined with gradient-boosted decision frameworks provide superior generalization, reduced misclassification, and strong robustness against acoustic variability. The system also proved highly reliable in noisy industrial environments, with the Hist GB model showing minimal confusion across all four classes. The developed GUI provides a user-friendly interface, enabling seamless dataset loading, feature extraction, model training, and real-time audio fault prediction. The research demonstrates a substantial performance enhancement compared to traditional machine learning approaches and establishes a strong foundation for deploying scalable acoustic monitoring solutions in industrial settings.

References

- [1] Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R.X. Deep learning, and its applications to machine health monitoring: A survey. *IEEE Trans. Neural Netw. Learn. Syst.* 2021, 14, 213–237.
- [2] Xiao, D.; Qin, C.; Yu, H.; Huang, Y.; Liu, C. Unsupervised deep representation learning for motor fault diagnosis by mutual information maximization. *J. Intell. Manuf.* 2021, 32, 377–391.
- [3] Zhou, F.; Sun, T.; Hu, X.; Wang, T.; Wen, C. A sparse denoising deep neural network for improving fault diagnosis performance. *Signal Image Video Process.* 2021, 15, 1889–1898.
- [4] Vasagam, M., Kumar, A., & Garg, A. (2026). Learning Execution Plan Embeddings for Multi-Dimensional Query Resource Prediction. *IEEE Access*.
- [5] Patyrykin, K. (2025). CANCEL CULTURE PROBLEM. *Lex Localis: Journal of Local Self-Government*, 23.
- [6] Yang, Y.; Haque, M.M.M.; Bai, D.; Tang, W. Fault Diagnosis of Electric Motors Using Deep Learning Algorithms and Its Application: A Review. *Energies* 2021, 14, 7017. <https://doi.org/10.3390/en14217017>
- [7] Kalae, U. K. (2023). Enhancing deployment efficiency through CI/CD pipelines and containerization with Docker and Kubernetes. *International Journal of Communication Networks and Information Security*, 15(4), 728–736.
- [8] Khaneghah, M.Z.; Alzayed, M.; Chaoui, H. Fault Detection and Diagnosis of the Electric Motor Drive and Battery System of Electric Vehicles. *Machines* 2023, 11, 713. <https://doi.org/10.3390/machines11070713>
- [9] Poojari, R. (2024). Empirical Analysis of Context Window Enhancement Methods in Retrieval-Augmented Generation Models. *Journal of Computational Analysis and Applications*, 33(2).
- [10] Purmani, S. S. R. (2025). Streamlining IT operations and service management with agile frameworks. *European Journal of Advances in Engineering and Technology*, 12(4), 76–81
- [11] Li, Q.-Y.; Wong, P.-K.; Vong, C.-M.; Fei, K.; Chan, I.-N. A Novel Electric Motor Fault Diagnosis by Using a Convolutional Neural Network, Normalized Thermal Images and Few-Shot Learning. *Electronics* 2024, 13, 108. <https://doi.org/10.3390/electronics13010108>

-
- [12] Huang, K.; Chen, Q. Motor Fault Diagnosis and Predictive Maintenance Based on a Fine-Tuned Qwen2.5-7B Model. *Processes* 2025, 13, 2051. <https://doi.org/10.3390/pr13072051>
- [13] Tao, Y.; Wang, X.; Zhang, L.; Bao, X.; Xue, H.; Yue, H.; Feng, H.; Yang, D. Fault Diagnosis of In-Wheel Motors Used in Electric Vehicles: State of the Art, Challenges, and Future Directions. *Machines* 2025, 13, 711. <https://doi.org/10.3390/machines13080711>
- [14] Kumar, P.; Prince; Sinha, A.K.; Kim, H.S. Electric Vehicle Motor Fault Detection with Improved Recurrent 1D Convolutional Neural Network. *Mathematics* 2024, 12, 3012. <https://doi.org/10.3390/math12193012>
- [15] Sobhi, S.; Reshadi, M.; Zarft, N.; Terheide, A.; Dick, S. Condition Monitoring and Fault Detection in Small Induction Motors Using Machine Learning Algorithms. *Information* 2023, 14, 329. <https://doi.org/10.3390/info14060329>
- [16] Soresini, F.; Barri, D.; Cazzaniga, I.; Ballo, F.M.; Mastinu, G.; Gobbi, M. Artificial Intelligence for Fault Detection of Automotive Electric Motors. *Machines* 2025, 13, 457. <https://doi.org/10.3390/machines13060457>.